DOCUMENT RESUME

ED 263 376 CE 042 889

AUTHOR MacFarland, Thomas W.
TITLE A Defense of Criterion-Referenced Evaluation for Vocational Education.
PUB DATE Dec 85
NOTE 6p.
PUB TYPE Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; *Criterion Referenced Tests; *Norm Referenced Tests; Secondary Education; *Student Evaluation; *Test Reliability; *Test Validity; *Vocational Education
IDENTIFIERS *Behaviorism

ABSTRACT

Criterion-referenced evaluation (CRE) describes achievement in performance terms, whereas norm-referenced evaluation (NRE) compares the performance of one individual to that of others with respect to a given evaluation instrument. Vocational educators who base their programs on behaviorism commonly evaluate student performance from a CRE perspective. When deciding whether to use CRE or NRE to evaluate a given instance of learning, vocational instructors must examine their purposes in undertaking the given evaluation. In situations when an individual's performance must be measured in terms of comparison with a standardized sample, NRE is the obvious choice. However, vocational educators must also realize the value of CRE in assessing the effectiveness of given instructional treatments in achieving specific instructional objectives. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED263376

A DEFENSE OF CRITERION-REFERENCED EVALUATION FOR VOCATIONAL EDUCATION

by

Thomas W. MacFarland

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

T MacFarland

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CE 042889

Nova University Center for Higher Education
Fort Lauderdale, FL 33314

December, 1985

INTRODUCTION

The term criterion-referenced evaluation (CRE) was first proposed by Glaser in 1963 (Anastasi, 1982:94) and is used, according to Mehrens and Lehmann (1978:692)

> to provide information on the specific knowledge or skills possessed by a student. . . . [with] meaning in terms of what the student knows or can do, rather than in their relation to the scores made by some external reference group.

CRE, therefore, "describes achievement in performance terms (Gronlund, 1981:416) and is in contrast to norm-referenced evaluation (NRE) where "an individual's score is interpreted by comparing it with the scores obtained by others on the same test" (Anastasi, 1982:94).

Skinner purported direct assessment of target behavior since trait names such as aggressive, lazy, and retarded fail to refer to any unit of behavior suitable for study through functional analysis (Bower and Hilgard, 1981: 198). Regarding Skinner, Mehrens and Lehmann (1978:51-52) attributed CRE to behaviorism and the

> emphases on behavioral objectives, the sequencing and individualization of instruction, the development of programmed materials . . . and a belief that norm referencing promotes unhealthy competition and is injurious to low-scoring-students' self-concepts.

Vocational instructors who use behaviorism commonly evaluate student performance from a CRE perspective. To better understand the appropriate uses of CRE, it is necessary to grasp the context of reliability and validity from a CRE position.

RELIABILITY FROM A CRITERION-REFERENCED POSITION

Reliability is defined by Mehrens and Lehmann (1978: 700) as "the extent to which a test is consistent in measuring whatever it does measure." Anastasi (1982:118) identified the following as sources of error variance in relation to their respective reliability coefficients: time sampling (test-retest); content sampling (immediate alternate-form); time sampling and content sampling (delayed alternate-form); content sampling (split-half); content sampling and content heterogeneity (Kuder-Richardson and coefficient Alpha); and interscorer differences (scorer).

In view of the determination of reliability, Popham and Husek (1973:134) have identified how, because there may be minimal score variability, the classical procedures that are used to assess reliability are not appropriate. In contrast to NRE, a CRE test may have not only zero internal consistency, but the test may also have negative internal consistency and still be a good test. Anastasi (1982:130) addressed the issue of CRE reliability and stated that CRE test data can be used "provided they are applied to the appropriate data. . . . the test scores of persons who have had different amounts of prior training."

VALIDITY FROM A CRITERION-REFERENCED POSITION

Validity is defined by Mehrens and Lehmann (1978:109) as "the degree to which a test is capable of achieving

certain aims" and is further described as "the single most important aspect of a test." Three general types of validity have been identified. Anastasi (1982:131) defines content validity as "the systematic examination of the test content to determine whether it covers a representative sample of the behavior domain to be measured." Criterion-related validity is identified by Mehrens and Lehmann (1978:703-704) as "the extent to which scores on the test are in agreement with (concurrent validity) or predict (predictive validity) some given criterion measure." Construct validity is described by Anastasi (1982:144) as the extent theoretical constructs or traits such as intelligence or mechanical comprehension are measured. In regard to CRE validity, Anastasi (1982:140) stated, "the most satisfactory type of criterion measure is that based on follow-up records of actual job performance."

CONCLUSION AND RECOMMENDATIONS

Vocational instructors must realize that the decision to use either CRE or NRE depends on the purpose of the evaluation process. CRE, which is used to "assess the effectiveness of given instructional treatments in achieving specific instructional objectives" (Isaac and Michael, 1981:110), can be as equally appropriate as NRE which must consider "the test performance of the standardized sample" (Anastasi, 1982: 65). Although NRE has its appropriate applications, vocational instructors must also recognize the appropriate uses of CRE as they assess achievement of predetermined criteria.

REFERENCES

Anastasi, Ann. Psychological Testing. 5th ed. New York: MacMillan Publishing Company, Inc., 1982.

Bower, Gordon H., and Ernest R. Hilgard. Theories of Learning. 5th ed. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1981.

Gronlund, Norman E. Measurement and Evaluation in Teaching. 4th ed. MacMillan Publishing Company, Inc., 1981.

Isaac, Stephen, and William B. Michael. Handbook in Research and Evaluation. 2nd ed. San Diego: EdITS Publishers, 1981.

Mehrens, William A., and Irvin J. Lehmann. Measurement and Evaluation in Education and Psychology. 2nd ed. New York: Holt, Rinehart and Winston, 1978.

Popham, W. James, and T. R. Husek. "Implications of Criterion-Referenced Measurement." In Toward Instructional Accountability. Eds. John E. Roueche and Barton R. Herrscher. Palo Alto, California: Westinghouse Learning Press, 1973.